

Adaptive phase measurements for narrowband squeezed beams

Dominic W. Berry¹ and Howard M. Wiseman²

¹*Department of Physics, The University of Queensland, Brisbane, Queensland 4072, Australia**

²*Centre for Quantum Dynamics, School of Science,
Griffith University, Nathan, Brisbane, Queensland 4111, Australia†*

(Dated: February 1, 2008)

We have previously [Phys. Rev. A **65**, 043803 (2002)] analyzed adaptive measurements for estimating the continuously varying phase of a coherent beam, and a broadband squeezed beam. A real squeezed beam must have finite photon flux \mathcal{N} and hence can be significantly squeezed only over a limited frequency range. In this paper we analyze adaptive phase measurements of this type for a realistic model of a squeezed beam. We show that, provided it is possible to suitably choose the parameters of the beam, a mean-square phase uncertainty scaling as $(\mathcal{N}/\kappa)^{-5/8}$ is possible, where κ is the linewidth of the beam resulting from the fluctuating phase. This is an improvement over the $(\mathcal{N}/\kappa)^{-1/2}$ scaling found previously for coherent beams. In the experimentally realistic case where there is a limit on the maximum squeezing possible, the variance will be reduced below that for coherent beams, though the scaling is unchanged.

PACS numbers: 42.50.Dv, 42.50.Lc

I. INTRODUCTION

Optical phase measurements are a valuable means for high precision measurement of displacement, for example for gravitational wave detection [1]. Two types of phase measurement can be distinguished. The first is measuring the phase of a beam or pulse of light relative to that of a strong local oscillator, which is treated classically. The second is measuring the phase shift between the light beams in the two arms of an interferometer, both of which are treated quantum mechanically. In this paper we are concerned solely with the first type.

Phase measurements are most easily analyzed, and thus best understood, for a single mode pulse with a fixed phase. The optimal measurement scheme (for all commonly produced states) for such a single-shot measurement is the so-called canonical measurement [2]. For coherent states, where the quantum noise is independent of the quadrature, the canonical phase variance [2] asymptotically approaches $1/4\bar{n}$ in the large \bar{n} limit, where \bar{n} is the mean photon number. The $1/\bar{n}$ scaling is referred to as the standard quantum limit (SQL). On the other hand, for more general states where the noise is quadrature-dependent, the canonical phase variance may approach the Heisenberg limit $1.89/\bar{n}^2$ [3].

Unfortunately it is not possible to achieve canonical measurements with linear optical elements (unless one discards most of the measurement results [4]). One solution that has been developed is to use feedback. The field is combined with a local oscillator at a beam splitter, and partial results during the measurement are fed back to adjust the phase of the local oscillator for measuring the field in the next part of the pulse. This approach has

been extensively studied for single-shot measurements, and it has been shown that it can achieve scalings almost at the Heisenberg limit [5]. It has also been experimentally realized [6], verifying an improvement over non-adaptive (heterodyne) detection. However, because the experiment used coherent states, it was not possible to verify a scaling better than the SQL.

In practice it is easier to produce a continuous squeezed beam rather than a squeezed pulse of light. This motivates considering continuous, rather than single-shot, measurements. In the continuous case, if the phase to be measured is constant, then the variance will become arbitrarily small with time. To obtain a nontrivial result, it is necessary to consider a phase which varies in time. We wish to determine how accurately the measurement scheme estimates this varying phase. The simplest model for a varying phase is a Wiener process; that is, its rate of change is Gaussian white noise with intensity κ , giving rise to a Lorentzian lineshape of the beam with width κ .

Continuous measurements of this type were considered in Refs. [7, 8]. In Ref. [7], coherent beams and broadband squeezed beams were considered, and a simple method for filtering the data to obtain the phase estimate was used. A scaling law for the optimal variance (or mean-square error) of the squeezed beam was derived which showed an improvement over the coherent beam result. However, as was pointed out in Ref. [8], this analysis had two shortcomings (which tend to counteract each-other). First, it considered only the photon flux due to the coherent component of the beam, whereas in fact the photon flux from the broadband squeezing is strictly infinite. Second, its filtering technique ignored the fact that phase information may be obtained from the photocurrent *noise* [9]. For broadband squeezing the information from the noise is strictly infinite, and would allow the phase to be determined exactly, modulo π .

Ref. [8] adopted a more sophisticated (Bayesian) filtering of the state, but restricted its attention to coherent

*Electronic address: berry@physics.uq.edu.au

†Electronic address: H.Wiseman@griffith.edu.au

states. It found different results for $\mathcal{N} \lesssim \kappa$, though the asymptotic variance of $(1/2)\sqrt{\kappa/\mathcal{N}}$ for $\mathcal{N} \gg \kappa$ was the same as found in Ref. [7]. An asymptotic improvement by a factor of $1/\sqrt{2}$ over non-adaptive (heterodyne) detection was also confirmed.

In this paper we apply a Bayesian approach to phase measurements on a *narrowband* squeezed beam. This estimation procedure optimally uses the phase information obtained from both the mean field and from the noise. The narrowband squeezing ensures that the photon flux, and the rate of information accumulation in the photocurrent record, are finite. To obtain analytical results, we concentrate on the asymptotic regime. We find an approximate analytical expression for the optimal scaling of the phase variance as $(\mathcal{N}/\kappa)^{5/8}$, and confirm this via numerical simulations. Our analysis confirms that there is a scaling advantage over the SQL, although it is slightly less than we had previously thought [7].

The remainder of this paper is organized as follows. In Sec. II we give a simple explanation of the scaling, and explain why it is different from that in Ref. [7]. Then we give the Bayesian analysis of measurements with feedback in Sec. III, and of heterodyne measurements in Sec. IV. We revisit the scaling based on the Bayesian treatment in Sec. V, then give numerical results in Sec. VI. We conclude in Sec. VII, and give additional details of the derivations in the appendices.

II. SIMPLE EXPLANATION OF SCALING

In this section we give a simple explanation to predict the scaling of the phase variance under continuous adaptive measurements. First we discuss the simple case of a coherent beam, then we proceed to the case of squeezing.

A. Coherent states

The configuration considered is as in Fig. 1, although for the moment everything to the left of the dashed line should be ignored. The coherent signal beam has photon flux (i.e. mean number of photons per unit time) \mathcal{N} . A phase shift $\theta(t)$ (to be estimated by the experimenter) is imposed on the signal, and a known phase shift $\Phi(t)$ may be imposed by the experimenter on the local oscillator. These are combined at a beam splitter, and the difference photocurrent $I(t)$ gives a measurement of a quadrature of the signal. Feedback may be used to determine the phase $\Phi(t)$ for the local oscillator using a signal processor to filter the signal $I(t)$. The estimate of $\theta(t)$ is also determined based on $I(t)$.

For fixed system phase and a coherent beam, the variance for adaptive measurements over time interval Δt is

$$\sigma^2 \approx \frac{1}{4\Delta t \mathcal{N}}. \quad (2.1)$$

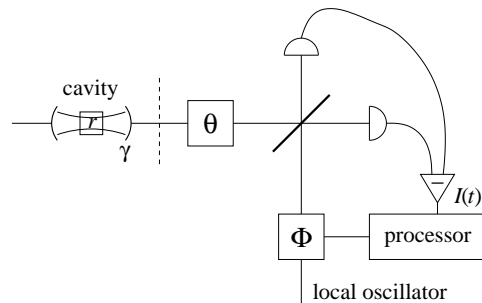


FIG. 1: The model of the experiment. The phase shift θ is imposed on the continuous beam, and a phase shift Φ is imposed on the local oscillator. These beams are incident on a 50:50 beam splitter, and the difference photocurrent $I(t)$ is determined. The processor then adjusts Φ based on $I(t)$. The simplest case is when the signal beam is in a coherent state, but for more accurate phase estimation a squeezed beam, produced by the apparatus to the left of the dashed vertical line, is used. This consists of a parametric down-converter characterized by intensity damping rate γ and $\chi^{(2)}$ nonlinearity parameterized by r .

This is simply the standard result for adaptive measurements on a coherent field with a mean photon number of $\Delta t \mathcal{N}$ [5]. If there is an estimate of the phase at time t , then the phase estimate at time $t + \Delta t$ may be taken to be a weighted average of the estimate at time t and the estimate obtained from data in the time interval Δt . The variance at time $t + \Delta t$ will then satisfy

$$\sigma_{t+\Delta t}^2 \approx \frac{1}{1/\sigma_t^2 + 4\Delta t \mathcal{N}}. \quad (2.2)$$

Now we assume that the phase fluctuations on the signal are white:

$$\dot{\theta} = \sqrt{\kappa} \zeta, \quad (2.3)$$

where ζ is Gaussian white noise satisfying $\langle \zeta(t)\zeta(t') \rangle = \delta(t-t')$. Taking account of these fluctuations, we expect the variance at time $t + \Delta t$ to be

$$\sigma_{t+\Delta t}^2 \approx \frac{1}{1/\sigma_t^2 + 4\Delta t \mathcal{N}} + \kappa \Delta t. \quad (2.4)$$

Provided $\kappa \Delta t \ll \sigma_t^2 \ll 1/(4\Delta t \mathcal{N})$, the total change in the inverse variance is

$$\Delta \left(\frac{1}{\sigma^2} \right) \approx -\Delta t \frac{\kappa}{\sigma^4} + 4\Delta t \mathcal{N}. \quad (2.5)$$

The steady-state variance is therefore

$$\sigma^2 \approx \frac{1}{2} \sqrt{\kappa/\mathcal{N}}. \quad (2.6)$$

This is just what was obtained in Ref. [7].

In the case of heterodyne measurements, the variance for measurements over time interval Δt is twice that in

Eq. (2.1) [5]. Then the last term in Eq. (2.5) is $2\Delta t\mathcal{N}$, and the steady state variance is

$$\sigma^2 \approx \frac{1}{\sqrt{2}} \sqrt{\kappa/\mathcal{N}}. \quad (2.7)$$

Thus the adaptive measurements give a $1/\sqrt{2}$ reduction in the variance over heterodyne measurements.

B. Squeezed states

We may use a similar method in the case of squeezed states. In this case we have a number of new features to the model, indicated in Fig. 1 by the apparatus to the left of the dashed line. The squeezed beam is produced by a cavity with decay constant γ . The beam has squeezing parameter r and coherent amplitude E . The flux is given by [10]

$$\mathcal{N} = \frac{E^2}{4} + \frac{\gamma}{2} \sinh^2 r. \quad (2.8)$$

For squeezing which is not too large, the phase variance for adaptive measurements over time interval Δt is, using the simple estimation technique of Ref. [7],

$$\sigma^2 \approx \frac{1}{e^{2r}\Delta t E^2}. \quad (2.9)$$

That is, the variance is reduced by a factor of e^{2r} from what would be obtained for a coherent state. The equivalent of Eq. (2.5) is then

$$\Delta \left(\frac{1}{\sigma^2} \right) \approx -\Delta t \frac{\kappa}{\sigma^4} + e^{2r} \Delta t E^2. \quad (2.10)$$

Provided the dominant contribution to the flux is the coherent component, the steady-state variance is

$$\sigma^2 \approx \frac{1}{2} e^{-r} \sqrt{\kappa/\mathcal{N}}. \quad (2.11)$$

There is a limit on how large the squeezing can be before this approximation fails. In Ref. [7] the limitation considered was the error in the feedback phase, which causes the measurement to not be exactly on the squeezed quadrature. This limitation gives an overall scaling of $(\kappa/\mathcal{N})^{2/3}$.

Here we consider the additional limitation due to the finite squeezing bandwidth. For measurements performed over a time scale shorter than $1/\gamma$, no squeezing will be observed. In fact, it is necessary to perform measurements over a time scale of order e^r/γ before squeezing is observed. For e^r large, this time is approximately the reciprocal of the decay constant for one of the cavity quadratures [see Eq. (3.2) in the following section]. However, it is not obvious why this decay constant is important, because it relates to the antisqueezed quadrature. We give a more thorough explanation of this time scale in the following section.

The phase information used to obtain the phase estimate will be from a finite time interval. For times before $t - \sigma^2/\kappa$, the system phase will differ from the current system phase by an amount comparable with the phase uncertainty. Therefore, the majority of the phase information must be taken from a time interval of length approximately σ^2/κ . We use χ to denote the inverse of this time interval, so $\chi \sim \kappa/\sigma^2$.

In order to observe squeezing, this time interval must be longer than the time scale e^r/γ , so we require $\sigma^2/\kappa > e^r/\gamma$. Here we are concerned with scaling, so we ignore multiplicative constants in the rest of this section. In order to obtain the scaling we also need to consider the limitation due to the squeezing contribution to the flux. From Eq. (2.8), we have $\mathcal{N} > \gamma e^{2r}$, so

$$\sigma^2 > \frac{\kappa}{\mathcal{N}} e^{3r}. \quad (2.12)$$

In order to obtain the minimum phase variance, we need to take the squeezing r to be as large as possible consistent with the Eqs. (2.11) and (2.12). That gives $e^r \sim (\mathcal{N}/\kappa)^{1/8}$, so

$$\sigma^2 \sim \left(\frac{\kappa}{\mathcal{N}} \right)^{5/8}. \quad (2.13)$$

Note that if we had used the time scale $1/\gamma$, rather than e^r/γ , then we would have obtained the inequality $\sigma^2/\kappa > 1/\gamma$. Then Eq. (2.12) would become $\sigma^2 > e^{2r}\kappa/\mathcal{N}$, and the maximum squeezing would be $e^r \sim (\mathcal{N}/\kappa)^{1/6}$. This then would give the variance scaling as $\sigma^2 \sim (\kappa/\mathcal{N})^{2/3}$, which is identical to that in Ref. [7]. Thus we can see that the extra e^r is essential to obtaining the different scaling here. This scaling will be shown to be correct numerically in Sec. VI.

In practice it is not possible to achieve arbitrary squeezing; typically the maximum value of e^{2r} achieved is about 2 [14]. In that case, we simply have the variance given in Eq. (2.11), with r taken to be the maximum experimentally achievable value. That is, the scaling is the same as for coherent states, but the variance is reduced by a factor of e^{-r} .

Note that we have the limitations on γ (ignoring constant factors)

$$e^{2r} \sqrt{\frac{\mathcal{N}}{\kappa}} < \frac{\gamma}{\kappa} < e^{-2r} \frac{\mathcal{N}}{\kappa}. \quad (2.14)$$

These limitations may be satisfied provided $\mathcal{N}/\kappa > e^{8r}$; for $e^{2r} \sim 2$ this limit is about 16. Experimentally there is only limited control over the value of γ . However, for larger \mathcal{N}/κ there is a wide range of values for which a reduction in the phase variance should be observed.

III. ADAPTIVE MEASUREMENTS

Now we give the detailed Bayesian analysis of the phase estimates for narrowband squeezing. We will use this to

justify some of the steps used in the previous section, as well as to estimate the additional phase information that may be obtained from the noise. A continuous squeezed beam produced by a cavity may be modeled by the operator equations in the Heisenberg picture [12]

$$\frac{d\hat{x}}{dt} = -\hat{x}\gamma(1+\varepsilon)/2 + \sqrt{\gamma}\hat{\xi} \quad (3.1)$$

$$\frac{d\hat{y}}{dt} = -\hat{y}\gamma(1-\varepsilon)/2 + \sqrt{\gamma}\hat{\eta} \quad (3.2)$$

$$\hat{I} = \cos(\Phi - \theta)[\sqrt{\gamma}\hat{x} - \hat{\xi}] + \sin(\Phi - \theta)[\sqrt{\gamma}\hat{y} + E - \hat{\eta}]. \quad (3.3)$$

These equations are equivalent to Eqs. (4.48) and (4.54) in Ref. [12]. The quantities \hat{x} and \hat{y} are the two quadratures of the cavity field, $\hat{\xi}$ and $\hat{\eta}$ are the quadrature noise operators, γ is the cavity decay constant (equivalent to κ_1 in [12]), and ε is a constant related to the usual squeezing parameter by

$$e^r = \frac{1+\varepsilon}{1-\varepsilon}. \quad (3.4)$$

\hat{I} is the output quadrature at angle $\Phi - \theta$; the photocurrent measured corresponds to the measured value of this operator. Note that the squeezed quadrature here is x (rather than y , as in [12]). Also we have added the displacement E , so we obtain a squeezed coherent field rather than a squeezed vacuum.

These equations may be solved by using the Wigner distribution, and replacing the quadrature operators with the corresponding quadrature variables for the Wigner distribution. Also the output quadrature \hat{I} is replaced with the detected photocurrent I .

$$\dot{x} = -x\gamma(1+\varepsilon)/2 + \sqrt{\gamma}\xi \quad (3.5)$$

$$\dot{y} = -y\gamma(1-\varepsilon)/2 + \sqrt{\gamma}\eta \quad (3.6)$$

$$I = \cos(\Phi - \theta)[\sqrt{\gamma}x - \xi] + \sin(\Phi - \theta)[\sqrt{\gamma}y + E - \eta]. \quad (3.7)$$

Here ξ and η are Gaussian increments satisfying $\langle \xi(t)\xi(t') \rangle = \langle \eta(t)\eta(t') \rangle = \delta(t-t')$.

In order to apply the Bayesian approach, we assume for the moment that both Φ and θ are constant. The results obtained using these assumptions should also be accurate for cases where these phases only change by a small amount over the time interval considered. We discretize the equations to give

$$\Delta x = -x\Delta t\gamma(1+\varepsilon)/2 + \tilde{\xi}\sqrt{\gamma\Delta t} \quad (3.8)$$

$$\Delta y = -y\Delta t\gamma(1-\varepsilon)/2 + \tilde{\eta}\sqrt{\gamma\Delta t} \quad (3.9)$$

$$I\Delta t = \{\cos(\Phi - \theta)[\sqrt{\gamma\Delta t}x - \tilde{\xi}] + \sin(\Phi - \theta)[\sqrt{\gamma\Delta t}(y + E') - \tilde{\eta}]\}\sqrt{\Delta t}. \quad (3.10)$$

Here $\sqrt{\gamma}E' = E$, and we have used tildes to indicate that the stochastic increments have been replaced with Gaussian random variables (with mean 0 and variance 1).

At all times the experimenter's knowledge can be represented by a probability distribution for x , y , and θ which is Gaussian for x and y :

$$P(\vec{x}, \theta) = P(\theta) \frac{\sqrt{\det G(\theta)}}{2\pi} e^{-\frac{1}{2}[\vec{x} - \bar{x}(\theta)]^T G(\theta) [\vec{x} - \bar{x}(\theta)]} = P(\theta)P(\vec{x}|\theta), \quad (3.11)$$

where $\vec{x} = (x, y)^T$ and $\bar{x}(\theta)$ is the mean of \vec{x} . $P(\theta)$ gives the correct probability for θ averaging over x and y , and the inverse of $G(\theta)$ is the covariance matrix for \vec{x} for a given θ .

A. Phase information

First we consider the update to the probability distribution due to the information from the measurement. At each time step we update the probability distribution using Bayes' rule [8]:

$$P(\vec{x}, \theta) \rightarrow P(\vec{x}, \theta) \times P(I|\vec{x}, \theta). \quad (3.12)$$

Here the constant factor $1/P(I)$ is omitted, because we can normalize at the end of the calculation. The probability $P(I|\vec{x}, \theta)$ is given by

$$\begin{aligned} P(I|\vec{x}, \theta) &= \sqrt{\frac{\Delta t}{2\pi}} \exp[-\Delta t\{I - \sqrt{\gamma}[\cos(\Phi - \theta)x \\ &\quad + \sin(\Phi - \theta)(y + E')]\}^2/2] \\ &= \sqrt{\frac{\Delta t}{2\pi}} \exp\left[-\frac{\gamma\Delta t}{2}(\vec{x} - B)^T A A^T (\vec{x} - B)\right], \end{aligned} \quad (3.13)$$

where

$$A = \begin{bmatrix} c \\ s \end{bmatrix}, \quad B = \frac{\Delta I}{2\sqrt{\gamma}} \begin{bmatrix} 1/c \\ 1/s \end{bmatrix}, \quad (3.14)$$

where $c = \cos(\Phi - \theta)$, $s = \sin(\Phi - \theta)$ and $\Delta I = I - sE$.

We therefore obtain

$$\begin{aligned} P(I|\vec{x}, \theta)P(\vec{x}, \theta) &\propto \sqrt{\det G} \\ &\times \exp\left\{-\frac{1}{2}[(\vec{x} - \bar{x}^{(2)})^T G^{(2)}(\vec{x} - \bar{x}^{(2)}) \right. \\ &\quad \left. - (\bar{x}^{(2)})^T G^{(2)}\bar{x}^{(2)} + \bar{x}^T G\bar{x} + \gamma\Delta t B^T A A^T B]\right\}, \end{aligned} \quad (3.15)$$

where

$$G^{(2)} = G + \gamma\Delta t A A^T \quad (3.16)$$

$$\bar{x}^{(2)} = (G^{(2)})^{-1}(G\bar{x} + \gamma\Delta t A A^T B), \quad (3.17)$$

Hence the updated probability distribution for θ is

$$\begin{aligned} P^{(2)}(\theta) &\propto P(\theta) \sqrt{\frac{\det G}{\det G^{(2)}}} \exp\left\{-\frac{1}{2}[\bar{x}^T G\bar{x} \right. \\ &\quad \left. + \gamma\Delta t (B^T A)^2 - (\bar{x}^{(2)})^T G^{(2)}\bar{x}^{(2)}]\right\}. \end{aligned} \quad (3.18)$$

In the limit of small Δt we obtain the differential equations

$$dG = \gamma dt AA^T, \quad (3.19)$$

$$d\bar{x} = dt G^{-1} [\sqrt{\gamma} \Delta I A - \gamma AA^T \bar{x}], \quad (3.20)$$

$$d[\log P(\theta)] = K - \frac{dt}{2} (\Delta I - \sqrt{\gamma} A^T \bar{x})^2, \quad (3.21)$$

where K is a constant; K may be ignored, as it only changes the normalization.

B. Increments in x and y

Now we take account of the increments in x and y . Given the measurement result I , we have the restriction that

$$c\tilde{\xi} + s\tilde{\eta} = \sqrt{\Delta t} [\sqrt{\gamma}(cx + sy) - \Delta I]. \quad (3.22)$$

We define the new variable

$$\begin{aligned} \tilde{\mu} &= -(\tilde{\xi} - c\sqrt{\Delta t} [\sqrt{\gamma}(cx + sy) - \Delta I])/s \\ &= (\tilde{\eta} - s\sqrt{\Delta t} [\sqrt{\gamma}(cx + sy) - \Delta I])/c. \end{aligned} \quad (3.23)$$

This is simply a Gaussian random variable with mean zero and variance 1. In terms of this the new difference equations are

$$\begin{aligned} \Delta x &= -x\Delta t\gamma(1 + \varepsilon)/2 - c\Delta t[\sqrt{\gamma}\Delta I - \gamma(cx + sy)] \\ &\quad - s\tilde{\mu}\sqrt{\gamma\Delta t} \end{aligned} \quad (3.24)$$

$$\begin{aligned} \Delta y &= -y\Delta t\gamma(1 - \varepsilon)/2 - s\Delta t[\sqrt{\gamma}\Delta I - \gamma(cx + sy)] \\ &\quad + c\tilde{\mu}\sqrt{\gamma\Delta t}. \end{aligned} \quad (3.25)$$

The deterministic part of the increment gives the mapping $\bar{x} \mapsto C\bar{x} - \sqrt{\gamma}\Delta t\Delta I A$ with

$$\begin{aligned} C &= \mathbb{1} - \Delta t\gamma[(\mathbb{1} + \varepsilon\Sigma)/2 - AA^T], \\ \Sigma &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned} \quad (3.26)$$

The stochastic part increases the covariance matrix for \bar{x} according to $G \mapsto (G^{-1} + \gamma\Delta t\Pi)^{-1}$, where

$$\Pi = \begin{bmatrix} s^2 & -sc \\ -sc & c^2 \end{bmatrix}. \quad (3.27)$$

Overall, we update the covariance matrix and mean to

$$\begin{aligned} G^{(3)} &= [C(G^{(2)})^{-1}C + \gamma\Delta t\Pi]^{-1}, \\ \bar{x}^{(3)} &= C\bar{x}^{(2)} - \sqrt{\gamma}\Delta t\Delta I A. \end{aligned} \quad (3.28)$$

In the limit of infinitesimal Δt we again obtain differentials, which when added to those at the end of Sec. III A give

$$\begin{aligned} dG &= \gamma dt [AA^T + G + \varepsilon(\Sigma G + G\Sigma)/2 \\ &\quad - (AA^T G + GAA^T + G\Pi G)], \end{aligned} \quad (3.29)$$

$$\begin{aligned} d\bar{x} &= -\gamma dt (\mathbb{1} + \varepsilon\Sigma)\bar{x}/2 \\ &\quad + \sqrt{\gamma} dt (G^{-1} - \mathbb{1})A(\Delta I - \sqrt{\gamma}A^T\bar{x}). \end{aligned} \quad (3.30)$$

C. Solution

From Eq. (3.21), the probability distribution for the phase is obtained by integrating over $(\Delta I - \sqrt{\gamma}A^T\bar{x})^2$. From the reasoning given in Appendix A, we may estimate the final phase variance by determining the expectation value for a given phase Θ , and expanding to second order in $\theta - \Theta$. The probability distribution for I is determined based on Θ , and we determine the probability distribution for θ based on the measurement result I . We interpret Θ as the actual system phase, and θ as a dummy variable used for the probability distribution.

For fixed system phase and feedback phase, G will reach an equilibrium value. To estimate this equilibrium value, we rotate \bar{x} via

$$R = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad (3.31)$$

which is defined to simplify AA^T and Π . Using bars to denote the rotated values of variables, so for example $\bar{G}^{-1} = RG^{-1}R^\dagger$, we have from Eq. (3.29) at steady state

$$\begin{aligned} \bar{G}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \bar{G}^{-1} + \bar{G}^{-1} + \varepsilon(\bar{\Sigma}\bar{G}^{-1} + \bar{G}^{-1}\bar{\Sigma})/2 \\ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \bar{G}^{-1} + \bar{G}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned} \quad (3.32)$$

with

$$\bar{\Sigma} = \begin{bmatrix} c^2 - s^2 & -2sc \\ -2sc & s^2 - c^2 \end{bmatrix}. \quad (3.33)$$

Using the notation $a = (\bar{G}^{-1})_{00}$, $b = (\bar{G}^{-1})_{01}$, $d = (\bar{G}^{-1})_{11}$, $X = (1 - \varepsilon)/2 + s^2\varepsilon$, $Y = sc\varepsilon$, we have the simultaneous equations

$$a^2 = 2(aX + bY), \quad (3.34)$$

$$b^2 + 2(Xd - Yb) = 1, \quad (3.35)$$

$$ba = Y(a + d). \quad (3.36)$$

Solving for a^2 gives

$$a^2 = 2[X^2 + Y^2 + \sqrt{(X^2 + Y^2)^2 + Y^2}]. \quad (3.37)$$

Using that equilibrium value, the solution for $\sqrt{\gamma}A^T\bar{x}$ is (for large time, so initial conditions may be ignored)

$$\begin{aligned} \sqrt{\gamma}A^T\bar{x} &= \gamma A^T \int_0^t \exp\{\gamma[(\mathbb{1} + \varepsilon\Sigma)/2 \\ &\quad + (G^{-1} - \mathbb{1})AA^T](u - t)\} (G^{-1} - \mathbb{1})A\Delta I(u) du. \\ &= \gamma \text{Re} \left(\Omega \int_0^t du e^{\gamma\Lambda(u-t)} \Delta I(u) \right), \end{aligned} \quad (3.38)$$

where

$$\Omega = a - 1 + i(a - 2X)/\Delta, \quad (3.39)$$

$$2\Lambda = a + i\Delta, \quad (3.40)$$

with $\Delta^2/2 = \sqrt{(X^2 + Y^2)^2 + Y^2} - (X^2 + Y^2)$.

Expanding $(\Delta I - \sqrt{\gamma} A^T \bar{x})^2$ and taking the expectation value gives

$$\begin{aligned} \langle (\Delta I - \sqrt{\gamma} A^T \bar{x})^2 \rangle &= \langle \Delta I^2(t) \rangle \\ &- 2\gamma \text{Re} \left(\Omega \int_0^t du e^{\gamma \Lambda(u-t)} \langle \Delta I(t) \Delta I(u) \rangle \right) \\ &+ \frac{\gamma^2 |\Omega|^2}{2} \int_0^t du \int_0^t dv e^{\gamma[\Lambda(u-t) + \Lambda^*(v-t)]} \langle \Delta I(u) \Delta I(v) \rangle \\ &+ \frac{\gamma^2}{2} \text{Re} \left(\Omega^2 \int_0^t du \int_0^t dv e^{\gamma \Lambda(u+v-2t)} \langle \Delta I(u) \Delta I(v) \rangle \right). \end{aligned} \quad (3.41)$$

Determining $\langle I(u) I(v) \rangle$ gives

$$\begin{aligned} \langle I(u) I(v) \rangle &= (1 - S^2) \{ \gamma \langle x(u) x(v) \rangle - \sqrt{\gamma} [\langle x(u) \xi(v) \rangle \\ &+ \langle x(v) \xi(u) \rangle] + \langle \xi(u) \xi(v) \rangle \} \\ &+ S^2 \{ \gamma \langle y(u) y(v) \rangle + E^2 - \sqrt{\gamma} [\langle y(u) \eta(v) \rangle \\ &+ \langle y(v) \eta(u) \rangle] + \langle \eta(u) \eta(v) \rangle \}, \end{aligned} \quad (3.42)$$

where $S = \sin(\Phi - \Theta)$. The solutions for x and y are

$$x(t) = \sqrt{\gamma} \int_0^t e^{\gamma(1+\varepsilon)(v-t)/2} \xi dv, \quad (3.43)$$

$$y(t) = \sqrt{\gamma} \int_0^t e^{\gamma(1-\varepsilon)(v-t)/2} \eta dv. \quad (3.44)$$

Using this we obtain

$$\begin{aligned} \langle I(u) I(v) \rangle &= \delta(u - v) + S^2 E^2 + S^2 \frac{\gamma \varepsilon}{1 - \varepsilon} e^{-\gamma(1-\varepsilon)|u-v|/2} \\ &+ (S^2 - 1) \frac{\gamma \varepsilon}{1 + \varepsilon} e^{-\gamma(1+\varepsilon)|u-v|/2}, \end{aligned} \quad (3.45)$$

and

$$\begin{aligned} \langle \Delta I(u) \Delta I(v) \rangle &= \delta(u - v) + (S - s)^2 E^2 \\ &+ S^2 \frac{\gamma \varepsilon}{1 - \varepsilon} e^{-\gamma(1-\varepsilon)|u-v|/2} \\ &+ (S^2 - 1) \frac{\gamma \varepsilon}{1 + \varepsilon} e^{-\gamma(1+\varepsilon)|u-v|/2}. \end{aligned} \quad (3.46)$$

Using this result gives

$$\begin{aligned} \langle (\Delta I - \sqrt{\gamma} A^T \bar{x})^2 \rangle &= \delta(0) + \gamma f(r, \Phi - \Theta, \theta - \Theta) \\ &+ [1 - \text{Re}(\Omega/\Lambda)]^2 (S - s)^2 E^2, \end{aligned} \quad (3.47)$$

where $f(r, \Phi - \Theta, \theta - \Theta)$ is a function independent of γ and E ; the explicit form will be given below. The term on the second line of Eq. (3.47) may be identified with the phase information obtained from the mean field (as it is proportional to E^2), and is the most important contribution. Expanding to second order, $(S - s)^2 \approx \cos^2(\Phi - \Theta)(\theta - \Theta)^2$, while

$$[1 - \text{Re}(\Omega/\Lambda)]^2 = (c^2 e^{-2r} + s^2 e^{2r})^{-1}. \quad (3.48)$$

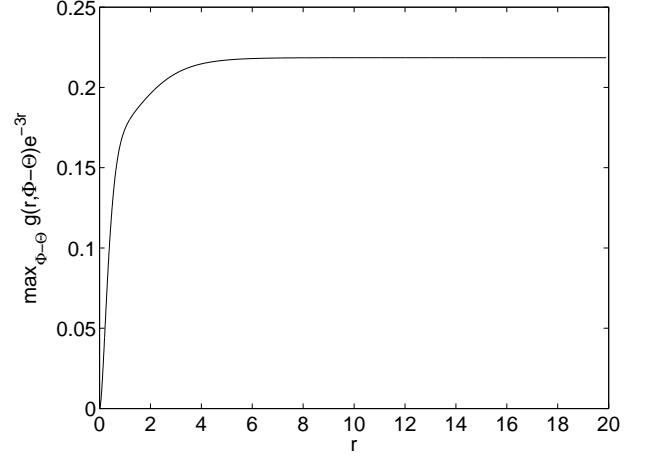


FIG. 2: The maximum of $g(r, \Phi - \Theta)$ divided by e^{3r} as a function of r .

In this expression we can replace s with S , because we only need this term to zeroth order in $\theta - \Theta$. We then have

$$\begin{aligned} \langle (\Delta I - \sqrt{\gamma} A^T \bar{x})^2 \rangle &= \frac{E^2 \cos^2(\Phi - \Theta)}{\Xi} (\theta - \Theta)^2 \\ &+ \gamma f(r, \Phi - \Theta, \theta - \Theta) + \delta(0) + O((\theta - \Theta)^3), \end{aligned} \quad (3.49)$$

where $\Xi = (1 - S^2)e^{-2r} + S^2 e^{2r}$.

The additional term $\gamma f(r, \Phi - \Theta, \theta - \Theta)$ gives the phase information obtained from the noise. Its explicit form is

$$\begin{aligned} f(r, \Phi - \Theta, \theta - \Theta) &= \text{Re} \left\{ \frac{|\Omega|^2}{2a} + \frac{\Omega^2}{4\Lambda} \right. \\ &+ S^2 \frac{\varepsilon}{1 - \varepsilon} \left[1 + \frac{\Omega^2/2\Lambda + |\Omega|^2/a - 2\Omega}{\Lambda + (1 - \varepsilon)/2} \right] \\ &\left. + (S^2 - 1) \frac{\varepsilon}{1 + \varepsilon} \left[1 + \frac{\Omega^2/2\Lambda + |\Omega|^2/a - 2\Omega}{\Lambda + (1 + \varepsilon)/2} \right] \right\}. \end{aligned} \quad (3.50)$$

Expanding this function to second order in $\theta - \Theta$ gives

$$\begin{aligned} f(r, \Phi - \Theta, \theta - \Theta) &= \text{const.} + g(r, \Phi - \Theta)(\theta - \Theta)^2 \\ &+ O((\theta - \Theta)^3), \end{aligned} \quad (3.51)$$

where

$$\begin{aligned} g(r, \Phi - \Theta) &= \frac{2 \sinh^2(r/2)}{\Xi a_\Theta} [(1 + \cosh r) \Delta_\Theta^2 \\ &+ 2 \sin^2(\Phi - \Theta) \cosh r \left(1 + \frac{\cosh r}{\sqrt{\Xi}} \right)], \end{aligned} \quad (3.52)$$

and a_Θ and Δ_Θ are equal to a and Δ for $\theta = \Theta$. Numerics indicate that $g(r, \Phi - \Theta) \leq e^{3r}/4$ (see Fig. 2); therefore this term may be ignored provided γe^{3r} is small compared to E^2/Ξ .

In Appendix A, the average over Θ is also taken. We take $\Phi - \Theta$ to be a constant, rather than taking Φ to be

independent of Θ . To obtain $\log P(\theta)$, we take the time integral of $(\Delta I - \sqrt{\gamma} A^T \bar{x})^2$ times $-1/2$. Therefore, for measurement over a time interval Δt we have

$$\frac{1}{\Delta t \sigma^2} \approx \frac{E^2 \cos^2(\Phi - \Theta)}{\Xi} + \gamma g(r, \Phi - \Theta). \quad (3.53)$$

The first term gives the phase information due to the coherent amplitude, whereas $\gamma g(r)$ gives the phase information due to the noise. Note that $\lim_{r \rightarrow 0} g(r) = 0$, so the phase information from the noise is zero for coherent states, as we expect.

An essential point is the time required for the system to come to equilibrium. The real part of the time constant $\gamma \Lambda$ is $a\gamma/2$. For small s , we have $a/2 \approx X$, so $\text{Re}(\Lambda) \approx 1/(e^r + 1)$. Therefore, the approximations made will only be accurate provided the measurement is made over a time period that is long compared to $(e^r + 1)/\gamma$. Ignoring the constant 1, this is the e^r/γ time scale used in the previous section.

IV. HETERODYNE MEASUREMENTS

We may derive similar results for the heterodyne case. The corresponding equations are

$$\dot{x} = -x\gamma(1 + \varepsilon)/2 + \sqrt{\gamma/2}(\xi^{(1)} + \xi^{(2)}) \quad (4.1)$$

$$\dot{y} = -y\gamma(1 - \varepsilon)/2 + \sqrt{\gamma/2}(\eta^{(1)} + \eta^{(2)}) \quad (4.2)$$

$$I_1 = \cos(\Phi - \Theta)[\sqrt{\gamma/2}x - \xi^{(1)}] + \sin(\Phi - \Theta)[\sqrt{\gamma/2}(y + E') - \eta^{(1)}] \quad (4.3)$$

$$I_2 = -\sin(\Phi - \Theta)[\sqrt{\gamma/2}x - \xi^{(2)}] + \cos(\Phi - \Theta)[\sqrt{\gamma/2}(y + E') - \eta^{(2)}]. \quad (4.4)$$

The quantities I_1 and I_2 are the two Fourier components of the photocurrents at the frequency at which the local oscillator is detuned from the system [13]. Because of that detuning, Φ here is arbitrary.

To see that the results do not depend on the value of Φ , note that we may represent the measurement results by the complex current $I = e^{i\Phi}(I_1 + iI_2)$ given by

$$I = e^{i\Theta} \{ \sqrt{\gamma/2} [x + i(y + E')] - \nu_1 - \nu_2 \}. \quad (4.5)$$

where $\nu_1 = (\xi^{(1)} + i\eta^{(2)} + \xi^{(2)} + i\eta^{(1)})/2$ and $\nu_2 = e^{2i(\Theta - \Phi)}(\xi^{(1)} + i\eta^{(2)} - \xi^{(2)} - i\eta^{(1)})/2$ are independent complex Gaussian random variables. Note also that

$$\dot{x} + i\dot{y} = -x\gamma(1 + \varepsilon)/2 - iy\gamma(1 - \varepsilon)/2 + \sqrt{2\gamma}\nu_1. \quad (4.6)$$

Following the derivation in the same way as above yields (see Appendix B)

$$\frac{1}{\Delta t \sigma^2} = \frac{E^2}{1 + e^{-2r}} + 2\gamma h(r), \quad (4.7)$$

where

$$h(r) = \cosh r - (\varepsilon^2 + 1)^{-1/2}. \quad (4.8)$$

As in the adaptive case, the first term gives the phase information due to the coherent amplitude, and $2\gamma h(r)$ gives the phase information due to the noise. Again, $\lim_{r \rightarrow 0} h(r) = 0$, so the phase information from the noise is zero for coherent states. For large r , $h(r)$ scales as e^r ; thus the phase information from the noise will be negligible unless the contribution to the photon flux due to the squeezing is dominant. In the derivation in Appendix B the $(e^r + 1)/\gamma$ time constant again appears, indicating that this is the appropriate time scale to observe squeezing in this case also.

V. SCALING REVISITED

We now use these results to give a more detailed derivation of the scaling for the phase variance. The derivations given in Secs. III and IV are based on fixed system phase and feedback phase (for the adaptive case). However, provided these phases vary by a small amount over some time interval $\Delta t > (e^r + 1)/\gamma$, these equations should still give an accurate estimate of the phase information over this time interval.

First we consider the adaptive case. Provided the feedback phase is sufficiently accurate, the \cos in Eq. (3.53) may be replaced with 1, and $\Xi \approx e^{-2r}$. In addition, provided $\gamma e^r \ll E^2$ (which will be true asymptotically unless the majority of the photon flux is from the squeezing), the second term which gives the phase information from the noise may be ignored. For continuous measurements, Eq. (3.53) gives the change in the inverse variance. Therefore we have

$$\frac{1}{\Delta t} \Delta \left(\frac{1}{\sigma^2} \right) \approx e^{2r} E^2. \quad (5.1)$$

Taking into account the system phase varying as $\dot{\Theta} = \sqrt{\kappa}\zeta$, we obtain the additional term $-\kappa/\sigma^4$ given in Eq. (2.10).

We may then follow the derivation given in Sec. II to obtain the scaling (2.13). There are two assumptions used in this derivation which we should justify. One assumption is that the phase information obtained from the noise is negligible. For the scaling obtained in Sec. II, $E^2/\Xi \sim \mathcal{N}e^{2r}$, whereas $\gamma e^{3r} < \mathcal{N}e^r$. Thus the assumption that little phase information is obtained from the noise is reasonable.

The other assumption is that $\Xi \sim e^{-2r}$. To check this condition, note that $\sin^2(\Phi - \Theta)$ may be made as small as order σ^2 via feedback. Therefore

$$\sin^2(\Phi - \Theta)e^{2r} \sim \left(\frac{\kappa}{\mathcal{N}} \right)^{5/8} \left(\frac{\mathcal{N}}{\kappa} \right)^{1/4} = \left(\frac{\kappa}{\mathcal{N}} \right)^{3/8} \quad (5.2)$$

and

$$e^{-2r} \sim \left(\frac{\kappa}{\mathcal{N}} \right)^{1/4} > \sin^2(\Phi - \Theta)e^{2r}. \quad (5.3)$$

Thus $\Xi \sim e^{-2r}$, as required.

For the case of heterodyne measurements, the equivalent of Eq. (2.10) is

$$\frac{1}{\Delta t} \Delta \left(\frac{1}{\sigma^2} \right) \approx -\frac{\kappa}{\sigma^4} + \frac{E^2}{1 + e^{-2r}}. \quad (5.4)$$

Solving for σ^2 gives

$$\sigma^2 \approx \frac{\sqrt{1 + e^{-2r}}}{2} \sqrt{\frac{\kappa}{\mathcal{N}}}. \quad (5.5)$$

Here we have used the approximation that $\mathcal{N} \approx E^2/4$, which simply means that there is little photon flux from the squeezing. If there is no limit to the squeezing, the asymptotic variance should be

$$\sigma^2 \approx \frac{1}{2} \sqrt{\frac{\kappa}{\mathcal{N}}}. \quad (5.6)$$

VI. NUMERICAL RESULTS

As it is predicted that the phase information from the noise should be negligible it is reasonable to consider a non-Bayesian method which does not use phase information from the noise. The phase estimation method used for the majority of the calculations was similar to that in Ref. [7]. The quantities A_t and B_t were calculated as

$$A_t = \int_{-\infty}^t e^{\chi(u-t)} e^{i\Phi} I(u) du, \quad (6.1)$$

$$B_t = - \int_{-\infty}^t e^{\chi(u-t)} e^{2i\Phi}. \quad (6.2)$$

Note that these quantities are not related to the A and B used in Sec. III. The χ used here is equivalent to that used in Sec. II B, in that the time scale over which previous measurement results are used is $1/\chi$.

A good phase estimate may be obtained as $\arg C_t$, where $C_t = A_t + \chi B_t A_t^*$ [7]. It was found that poor results were obtained if $\arg C_t$ was used in the feedback, so the feedback used was

$$\Phi(t) = \arg(C_t^{1-\delta} A_t^\delta) \quad (6.3)$$

for a suitably chosen δ .

In the calculations it is not necessary to independently vary \mathcal{N} and κ . We may scale the time by κ , so we obtain the dimensionless parameters \mathcal{N}/κ , γ/κ and χ/κ . The other parameters which we may vary are r and δ ; these are already dimensionless. We may predict scalings in terms of these dimensionless parameters:

$$\begin{aligned} \sigma^2 &\sim (\kappa/\mathcal{N})^{5/8}, & e^r &\sim (\mathcal{N}/\kappa)^{1/8}, \\ \frac{\gamma}{\kappa} &\sim (\mathcal{N}/\kappa)^{3/4}, & \frac{\chi}{\kappa} &\sim (\mathcal{N}/\kappa)^{5/8}. \end{aligned} \quad (6.4)$$

The scaling of χ/κ is predicted from the fact that we use phase information from a time interval $\sim \sigma^2/\kappa$.

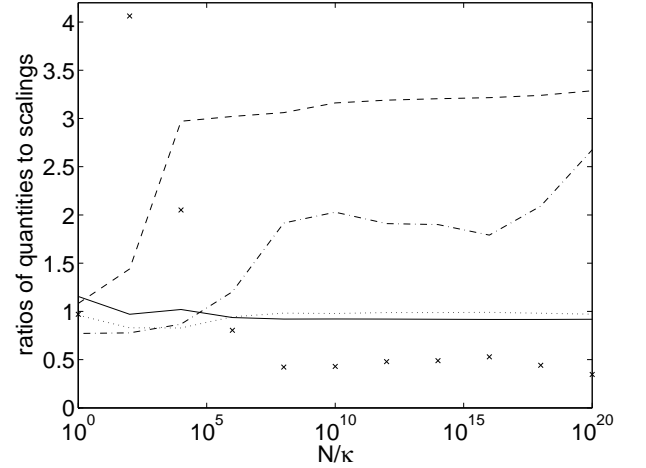


FIG. 3: The optimal values of various quantities as well as the variance for adaptive measurements with arbitrary squeezing. The values of $\sigma^2/(\mathcal{N}/\kappa)^{5/8}$ are shown as the solid line, $e^r/(\mathcal{N}/\kappa)^{1/8}$ is shown as the dotted line, $(\gamma/\kappa)/(\mathcal{N}/\kappa)^{3/4}$ is shown as the dashed line, $(\chi/\kappa)/(\mathcal{N}/\kappa)^{5/8}$ is shown as the dash-dotted line and $\delta/(\mathcal{N}/\kappa)^{1/4}$ is shown as the crosses.

Numerical calculations were performed for a range of values of \mathcal{N}/κ . For each value six alternative measurement schemes were considered:

1. Adaptive measurements with arbitrary squeezing;
2. Adaptive measurements with limited squeezing;
3. Adaptive measurements on coherent states;
4. Heterodyne measurements with arbitrary squeezing;
5. Heterodyne measurements with limited squeezing;
6. Heterodyne measurements on coherent states.

In each case the appropriate parameters were adjusted to minimize the variance. For all cases it was necessary to optimize over χ/κ ; for the adaptive measurement it was also necessary to optimize over δ . For the cases with squeezing the values of r and γ/κ were also optimized over. In the cases with limited squeezing there was the additional restriction that $e^{2r} \leq 2$. This limit corresponds to the typical maximum squeezing which may be achieved in the laboratory.

Results were obtained by integrating the system over a time interval of $130/\chi$, and determining the average phase variance for the data from time $30/\chi$ to $130/\chi$. A total of 2^{10} independent integrations were performed, so the total number of effectively independent samples, including those from the different integrations as well as those from different times within the integration, was approximately 10^5 .

The results for adaptive measurements on arbitrarily squeezed states are shown in Fig. 3. In this figure the various quantities were divided by the predicted scalings, in order to check that these scalings are correct. In each case we find that this ratio is of order 1, so the predicted scalings are correct. We also obtain the scaling for δ as

$$\delta \sim (\mathcal{N}/\kappa)^{1/4}. \quad (6.5)$$

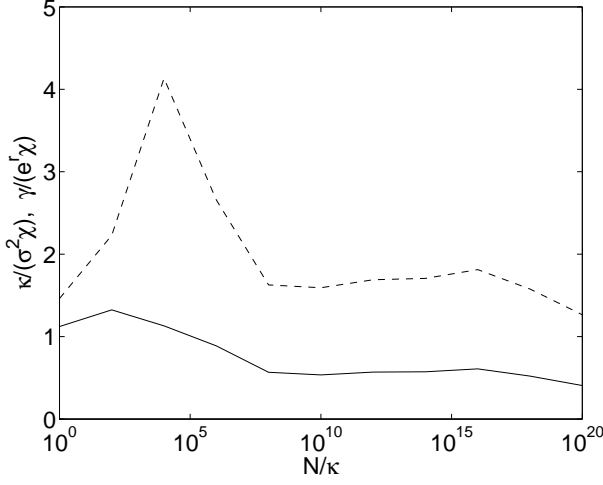


FIG. 4: The ratio of the value of κ/σ^2 to χ (continuous line), and the ratio of γ/e^r to χ (dashed line).

It does not appear to be possible to predict this scaling in the same way as the scalings for the other parameters given here.

The above theory predicts that χ should be approximately equal to κ/σ^2 in order to minimize the variance. Also, the time scale $1/\chi$ should be larger than the e^r/γ time scale required to obtain squeezing. To check these predictions, the ratio of κ/σ^2 to χ , as well as the ratio of γ/e^r to χ are shown in Fig. 4. From these results, the optimal value of χ is near κ/σ^2 , but there can be as much as a factor of two difference. We find that γ/e^r is larger than χ , as predicted. However, the difference is not great, and in some cases γ/e^r is only slightly larger than χ .

The phase variances for each of the six measurement schemes are shown in Fig. 5. In this figure the variances have been multiplied by $\sqrt{N/\kappa}$ in order to more clearly show the scaling constants. In each case except for adaptive measurements with arbitrarily squeezed states, it is clear that the variance is scaling as $\sqrt{\kappa/N}$.

It is straightforward to predict the asymptotic values of $\sigma^2 \sqrt{N/\kappa}$ under each of the measurement schemes using the results given in the preceding sections. These predictions are given in Table I, together with the asymptotic values estimated based on the numerical results. In each case, the predicted and calculated asymptotic values are within 1%.

Another issue, raised at the end of Sec. II B, is the range of possible γ values for which the benefit from squeezing may be observed. In that section we derived the upper and lower bounds on γ in Eq. (2.14). The range of values of γ/κ such that the phase variance was still within 10% of its minimum value was calculated numerically for adaptive phase measurements, and is shown in Fig. 6. (For the first two data points there was no lower bound found numerically, because the variance was close to that for a coherent state.)

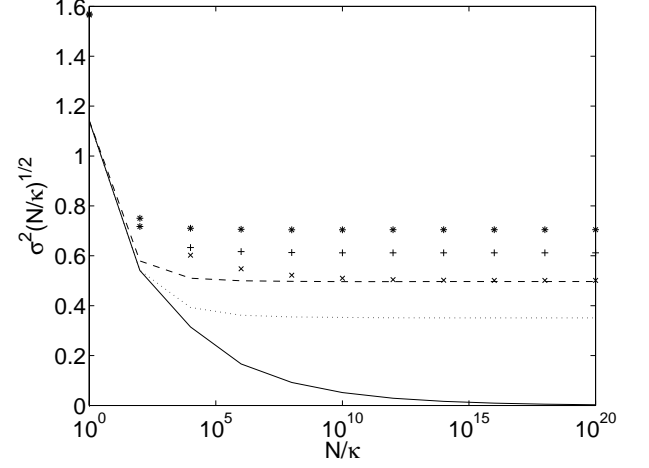


FIG. 5: The phase variance multiplied by $\sqrt{N/\kappa}$. The results for adaptive measurements on arbitrarily squeezed states, states limited to $e^{2r} \leq 2$, and coherent states are shown as the solid, dotted and dashed lines, respectively. The results for heterodyne measurements on arbitrarily squeezed states, states limited to $e^{2r} \leq 2$, and coherent states are shown as the crosses, plusses and asterisks, respectively.

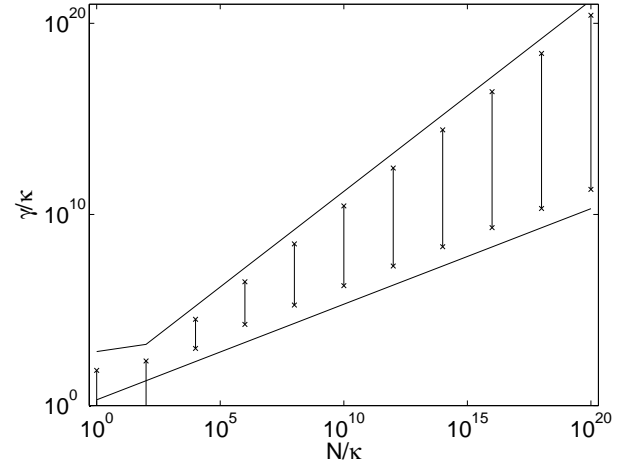


FIG. 6: The range of values of γ/κ for which it is possible to obtain phase variances within 10% of the minimum values with limited squeezing ($e^{2r} \leq 2$) and adaptive measurements. The ranges are shown by the vertical solid lines between the crosses, and the predicted upper and lower bounds $2N/(\kappa \sinh^2 r)$ and $e^{2r} \sqrt{N/\kappa}$ from Eq. (2.14) are shown by the diagonal solid lines.

The scaling of the upper and lower bounds found numerically is the same as the analytic bounds, though the range is slightly smaller. For the largest value of N/κ tested, there is a range of nine orders of magnitude for γ . This demonstrates that, even if it is not possible to adjust γ experimentally, there will be a wide range of values for which a reduction in the phase variance due to squeezing should be observed.

The last issue which we address in this section is that of using $\arg(C_t)$ rather than Bayesian phase estimates.

TABLE I: Asymptotic values of $\sigma^2 \sqrt{\mathcal{N}/\kappa}$ under each of the measurement schemes. The numerically estimated values are given first, and the analytic predictions are given in brackets.

| | adaptive | heterodyne |
|--------------------------------------|------------------------|------------------------|
| arbitrary squeezing | 0 (0) | 0.501 (1/2) |
| squeezing limited to $e^{2r} \leq 2$ | 0.351 ($1/\sqrt{8}$) | 0.612 ($\sqrt{3/8}$) |
| coherent states | 0.497 (1/2) | 0.705 ($1/\sqrt{2}$) |

To estimate the accuracy of the $\arg(C_t)$ phase estimates, calculations were also performed with Bayesian phase estimates. The system was integrated over a time period of $10^3/\chi$, and the phase variance was estimated by averaging from time $30/\chi$. The functions $P(\theta)$, $\bar{x}(\theta)$ and $G(\theta)$ were estimated by calculating them at 2000 values of θ . There is a complication when we take account of the variation in the system phase. If the Bayesian analysis is performed exactly, the distribution is no longer Gaussian in \bar{x} for given θ . It is not feasible to perform the calculation for the full distribution in \bar{x} and θ , so the distribution was approximated by a Gaussian in \bar{x} . The effect of the varying system phase on the distribution was approximated by simply adding a spread to the phase distribution $P(\theta)$.

For each set of data, the phase was estimated both via the Bayesian method and as $\arg(C_t)$. This allows accurate comparison of the relative variance. The ratio of the variance for the $\arg(C_t)$ phase estimates to that for the Bayesian phase estimates is shown in Fig. 7. There is only a small difference between the two variances; on average about 4%, and no more than 9%. In fact, the phase estimates obtained via $\arg(C_t)$ and the Bayesian method are quite close. The mean-square difference between these phase estimates is only about 3% of the total phase variance. This means that the qualitative results obtained for the $\arg(C_t)$ phase estimates should also hold when Bayesian phase estimates are used. The only difference is a reduction in the variance of a few percent.

The difference between the variances is likely due to the fact that the Bayesian estimate uses phase information from the noise, whereas the $\arg(C_t)$ estimate does not. Using the result (3.53) we may predict the ratio of the variances. This estimate is also shown in Fig. 7. There is not exact agreement with the numerically obtained ratio, but the estimate is close, particularly for the larger values of \mathcal{N}/κ .

VII. CONCLUSIONS

We have undertaken a thorough analysis of adaptive estimation of a continuously varying phase for narrow-band squeezed beams. The problem is characterized by photon flux \mathcal{N} , phase diffusion rate κ , squeezing bandwidth γ , and degree of squeezing e^{2r} (i.e. depth of squeezing of $1 - e^{-2r}$). If it is possible to achieve arbitrarily high squeezing, then adaptive phase measurements should give a variance scaling as $(\mathcal{N}/\kappa)^{-5/8}$, an improvement

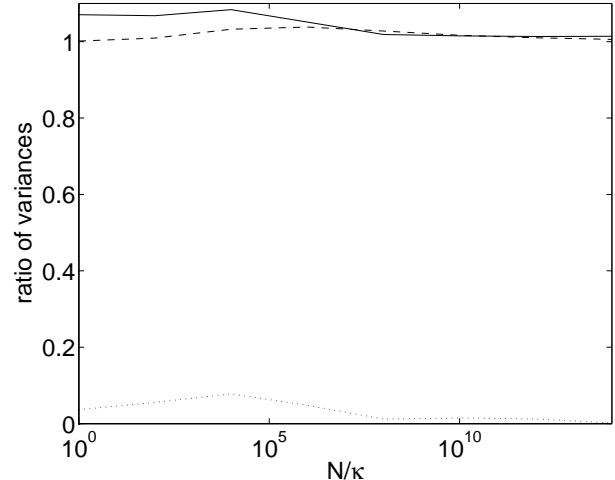


FIG. 7: The ratio of the phase variance obtained via $\arg(C_t)$ to the phase variance obtained via the Bayesian method (solid line). The estimated ratio, based on the difference being due to the phase information from the noise and using the maximum value of $g(r, \Phi - \Theta)$, is shown as the dashed line. The mean-square difference between the two phase estimates, as a ratio to the phase variance for Bayesian estimates, is shown as the dotted line.

over the $(\mathcal{N}/\kappa)^{-1/2}$ scaling for coherent beams. This variance is higher than that suggested by the broadband analysis in Ref. [7], which had a scaling of $(\mathcal{N}/\kappa)^{-2/3}$. The reason for this difference is that a time scale of e^r/γ is required before squeezing is observed. This means that the phase information must be obtained from a longer time interval than in the broadband case, and the system phase varies by a larger amount over this time interval. In the broadband case the limiting factor was the accuracy of the feedback.

It is somewhat surprising that the time scale required to observe squeezing is e^r/γ , rather than $1/\gamma$, which is what we would expect since γ is the squeezing bandwidth. The rate γ/e^r is the bandwidth of the anti-squeezed quadrature, but it is only through the full Bayesian analysis that one finds that this is the important rate. The difference of e^r is crucial, because the scaling of the phase variance would be $(\mathcal{N}/\kappa)^{-2/3}$ (as suggested by the broadband analysis) without it. The numerical calculations verify that the scaling is $(\mathcal{N}/\kappa)^{-5/8}$ rather than $(\mathcal{N}/\kappa)^{-2/3}$, demonstrating that the factor of e^r is correct.

We performed a Bayesian analysis of the phase, for

both adaptive and nonadaptive (heterodyne) measurement schemes. In each case the analysis yields a variance-reducing term proportional to E^2 , which may be interpreted as the phase information from the coherent amplitude, as well as a term proportional to γ , which corresponds to phase information from the squeezed noise. Except in cases where the dominant contribution to the photon flux is from the squeezing, the phase information from the noise is predicted to be negligible. This means that it is possible to obtain accurate phase estimates using a simplified method rather than the full Bayesian estimate.

In the experimentally realistic case where there is limited squeezing, the reduction in the phase variance is approximately e^r . Experimentally it is not possible to produce beams with arbitrary values of γ , though a limited amount of control is possible [14]. For larger values of \mathcal{N}/κ , the variance is insensitive to the exact value of γ , and similar variances are obtained for a wide range of values of γ . Therefore this issue is not expected to be a problem experimentally. A more complete analysis of the experimental feasibility of adaptive phase estimation on continuous squeezed beams will be given in a future work.

APPENDIX A: EXPECTED VARIANCE

The standard measure for phase uncertainty, used in most of the Refs. [5, 7, 8] is the Holevo variance [15]. It is defined as $|\langle e^{i\Delta\theta} \rangle|^{-2} - 1$, where $\Delta\theta$ is the error in the phase estimate; that is, the estimated phase minus the actual system phase. An alternative definition which avoids phase estimates with systematic error is $\text{Re}\langle e^{i\Delta\theta} \rangle^{-2} - 1$ [11]. The Holevo variance under this definition can be determined from [11]

$$\langle e^{i\Delta\theta} \rangle = \frac{1}{2\pi} \int dI \left| \int P(I|\theta) e^{i\theta} d\theta \right|. \quad (\text{A1})$$

Here I is used to indicate the entire measurement record. This may be alternatively written as

$$\begin{aligned} \langle e^{i\Delta\theta} \rangle &= \int P(I) dI \left| \int P(\theta|I) e^{i\theta} d\theta \right| \\ &= \frac{1}{2\pi} \int d\Theta \int P(I|\Theta) dI \left| \int P(\theta|I) e^{i\theta} d\theta \right|. \end{aligned} \quad (\text{A2})$$

The measurement record I can be expressed as a function of Θ as well as a record of random fluctuations ν which are independent of Θ . Therefore we may give the expression for $\langle e^{i\Delta\theta} \rangle$ as

$$\langle e^{i\Delta\theta} \rangle = \frac{1}{2\pi} \int d\Theta \int P(\nu) d\nu \left| \int P(\theta|\nu, \Theta) e^{i\theta} d\theta \right|. \quad (\text{A3})$$

Provided the measurement gives small variance, the only values of θ and Θ for which $P(\nu)P(\theta|\nu, \Theta)$ is not negligible are those for which θ is close to Θ . Thus it is

reasonable to expand $P(\theta|\nu, \Theta)$ as a series about Θ :

$$\langle e^{i\Delta\theta} \rangle \approx \frac{1}{2\pi} \int d\Theta \int P(\nu) d\nu \left| \int e^{a+b(\theta-\Theta)-c(\theta-\Theta)^2+i\theta} d\theta \right| \quad (\text{A4})$$

where a , b and c are functions of ν and Θ . Simplifying gives

$$\begin{aligned} \langle e^{i\Delta\theta} \rangle &\approx \frac{1}{2\pi} \int d\Theta \int P(\nu) e^{-1/(4c)} d\nu \\ &= \frac{1}{2\pi} \int d\Theta \langle e^{-1/(4c)} \rangle_{\Theta}. \end{aligned} \quad (\text{A5})$$

Provided the variance of c is small, it is possible to use the approximation

$$\langle e^{i\Delta\theta} \rangle \approx 1 - \frac{1}{8\pi} \int d\Theta (\langle c \rangle_{\Theta})^{-1}. \quad (\text{A6})$$

If $\langle c \rangle_{\Theta}$ is independent of Θ , $\langle e^{i\Delta\theta} \rangle \approx 1 - 1/(4\langle c \rangle)$, so the variance is approximately $1/2\langle c \rangle$.

APPENDIX B: HETERODYNE DERIVATION

The discretized version of the equations for heterodyne detection, (4.5) and (4.6) is

$$\begin{aligned} \Delta x + i\Delta y &= -\gamma\Delta t[x(1+\varepsilon)/2 + iy(1-\varepsilon)/2] + \sqrt{2\gamma\Delta t}\tilde{\nu}_1 \\ I &= e^{i\Theta}[\sqrt{\gamma/2}(x + iy + iE') - (\tilde{\nu}_1 + \tilde{\nu}_2)/\sqrt{\Delta t}]. \end{aligned} \quad (\text{B1})$$

Here the $\tilde{\nu}_1$, $\tilde{\nu}_2$ are complex Gaussian random variables with $\langle \tilde{\nu}_j \rangle = 0 = \langle \tilde{\nu}_i \tilde{\nu}_j \rangle$ but $\langle \tilde{\nu}_i^* \tilde{\nu}_j \rangle = \delta_{ij}$.

At each time step we update the probability $P(x, y, \theta)$, ignoring the constant factor $1/P(I)$, by

$$P(x, y, \theta) \rightarrow P(x, y, \theta) \times P(I|x, y, \theta). \quad (\text{B2})$$

The probability $P(I|x, y, \theta)$ is given by

$$\begin{aligned} P(I|x, y, \theta) &\propto \exp\left(-\frac{\Delta t}{2} \left| I - e^{i\theta} \sqrt{\gamma/2}(x + iy + iE') \right|^2\right) \\ &= \exp\left(-\frac{\gamma\Delta t}{4} [(x - I_x)^2 + (y - I_y)^2]\right), \end{aligned} \quad (\text{B3})$$

where $I_x + iI_y = \sqrt{2/\gamma} I e^{-i\theta} - iE'$. At all times the probability $P(x, y, \theta)$ is given by

$$P(x, y, \theta) = P(\theta)P(x|\theta)P(y|\theta), \quad (\text{B4})$$

where $P(x|\theta)$ and $P(y|\theta)$ are normal distributions for x and y . $P(\theta)$ gives the correct probability for θ averaging over x and y .

We therefore have

$$\begin{aligned} \frac{1}{\sigma_x^2} &\mapsto \frac{1}{\sigma_x^2} + \frac{\gamma\Delta t}{2}, & \frac{1}{\sigma_y^2} &\mapsto \frac{1}{\sigma_y^2} + \frac{\gamma\Delta t}{2}, \\ \bar{x} &\mapsto (\bar{x} + \sigma_x^2 I_x \gamma \Delta t / 2) / (1 + \sigma_x^2 \gamma \Delta t / 2), \\ \bar{y} &\mapsto (\bar{y} + \sigma_y^2 I_y \gamma \Delta t / 2) / (1 + \sigma_y^2 \gamma \Delta t / 2). \end{aligned} \quad (\text{B5})$$

In the infinitesimal limit,

$$\begin{aligned} d\sigma_x^2 &= -\sigma_x^4 \gamma dt/2, & d\sigma_y^2 &= -\sigma_y^4 \gamma dt/2, \\ d\bar{x} &= \sigma_x^2 (I_x - \bar{x}) \gamma dt/2, & d\bar{y} &= \sigma_y^2 (I_y - \bar{y}) \gamma dt/2. \end{aligned} \quad (\text{B6})$$

For the change in the probability distribution for θ we have, in the infinitesimal limit

$$P^{(2)}(\theta) \propto P(\theta) \exp \left\{ -\frac{\gamma dt}{4} [(I_x - \bar{x})^2 + (I_y - \bar{y})^2] \right\}. \quad (\text{B7})$$

Now we take account of the increments in x , y and θ . We have the restriction that

$$\tilde{\nu}_1 + \tilde{\nu}_2 = \sqrt{\Delta t} [\sqrt{\gamma/2} (x + iy + iE') - Ie^{-i\theta}]. \quad (\text{B8})$$

The mean of $\tilde{\nu}_1$ is therefore half of the RHS, and the second-moment is $1/2$, rather than 1 . Therefore, the difference equation may be expressed as

$$\Delta x + i\Delta y = -[(x - iy)\varepsilon + I_x + iI_y]\gamma\Delta t/2 + \tilde{\nu}\sqrt{\gamma\Delta t}, \quad (\text{B9})$$

where $\tilde{\nu}$ is a complex Gaussian random variable. Considering the deterministic part of the increment first, and taking the infinitesimal limit

$$\begin{aligned} d\sigma_x^2 &= -\sigma_x^2 \varepsilon \gamma dt, & d\sigma_y^2 &= \sigma_y^2 \varepsilon \gamma dt, \\ d\bar{x} &= -(I_x + \varepsilon \bar{x}) \gamma dt/2, & d\bar{y} &= -(I_y - \varepsilon \bar{y}) \gamma dt/2. \end{aligned} \quad (\text{B10})$$

To take account of the stochastic increments, we simply add the appropriate variances:

$$d\sigma_x^2 = \gamma dt/2, \quad d\sigma_y^2 = \gamma dt/2. \quad (\text{B11})$$

Overall we have the increments

$$\begin{aligned} d\sigma_x^2 &= [1 - 2\varepsilon\sigma_x^2 - \sigma_x^4] \gamma dt/2, \\ d\sigma_y^2 &= [1 + 2\varepsilon\sigma_y^2 - \sigma_y^4] \gamma dt/2. \end{aligned} \quad (\text{B12})$$

Therefore the steady state values are

$$\sigma_x^2 = \sqrt{\varepsilon^2 + 1} - \varepsilon, \quad \sigma_y^2 = \sqrt{\varepsilon^2 + 1} + \varepsilon. \quad (\text{B13})$$

The total increments in \bar{x} and \bar{y} are

$$\begin{aligned} d\bar{x} &= -[\bar{x}(\sigma_x^2 + \varepsilon) + (1 - \sigma_x^2)I_x] \gamma dt/2, \\ d\bar{y} &= -[\bar{y}(\sigma_y^2 - \varepsilon) + (1 - \sigma_y^2)I_y] \gamma dt/2. \end{aligned} \quad (\text{B14})$$

The solutions are

$$\begin{aligned} \bar{x} &= (\sigma_x^2 - 1)(\gamma/2) \int_0^t I_x e^{(\gamma/2)\sqrt{\varepsilon^2+1}(v-t)} dv, \\ \bar{y} &= (\sigma_y^2 - 1)(\gamma/2) \int_0^t I_y e^{(\gamma/2)\sqrt{\varepsilon^2+1}(v-t)} dv. \end{aligned} \quad (\text{B15})$$

Now $\langle I \rangle = ie^\Theta \sqrt{\gamma/2} E'$, so $\langle I_x + iI_y \rangle = i(e^{i(\Theta-\theta)} - 1)E'$, and

$$\begin{aligned} \langle \bar{x} \rangle &= \sin(\theta - \Theta) E' \frac{\sigma_x^2 - 1}{\sqrt{\varepsilon^2 + 1}}, \\ \langle \bar{y} \rangle &= [\cos(\theta - \Theta) - 1] E' \frac{\sigma_y^2 - 1}{\sqrt{\varepsilon^2 + 1}}. \end{aligned} \quad (\text{B16})$$

The solutions for x and y are

$$\begin{aligned} x(t) &= \sqrt{2\gamma} \int_0^t e^{\gamma(1+\varepsilon)(u-t)/2} \text{Re}(\nu_1) du, \\ y(t) &= \sqrt{2\gamma} \int_0^t e^{\gamma(1-\varepsilon)(u-t)/2} \text{Im}(\nu_1) du. \end{aligned} \quad (\text{B17})$$

In terms of these quantities,

$$\begin{aligned} I_x(t) &= x(t) \cos(\theta - \Theta) + (y(t) + E') \sin(\theta - \Theta) - \sqrt{2/\gamma} \{ \text{Re}(\nu_1 + \nu_2) \cos(\theta - \Theta) + \text{Im}(\nu_1 + \nu_2) \sin(\theta - \Theta) \} \\ I_y(t) &= -x(t) \sin(\theta - \Theta) + (y(t) + E') \cos(\theta - \Theta) - E' - \sqrt{2/\gamma} \{ \text{Im}(\nu_1 + \nu_2) \cos(\theta - \Theta) - \text{Re}(\nu_1 + \nu_2) \sin(\theta - \Theta) \}. \end{aligned} \quad (\text{B18})$$

Therefore

$$\begin{aligned} \langle I_x(u) I_x(v) \rangle &= \{ \langle x(u)x(v) \rangle - \sqrt{2/\gamma} \langle x(u) \text{Re}[\nu_1(v)] + x(v) \text{Re}[\nu_1(u)] \rangle \} \cos^2(\theta - \Theta) \\ &\quad + \{ \langle y(u)y(v) \rangle - \sqrt{2/\gamma} \langle y(u) \text{Im}[\nu_1(v)] + y(v) \text{Im}[\nu_1(u)] \rangle + (E')^2 \} \sin^2(\theta - \Theta) + (2/\gamma) \delta(u - v) \\ \langle I_y(u) I_y(v) \rangle &= \{ \langle x(u)x(v) \rangle - \sqrt{2/\gamma} \langle x(u) \text{Re}[\nu_1(v)] + x(v) \text{Re}[\nu_1(u)] \rangle \} \sin^2(\theta - \Theta) \\ &\quad + \{ \langle y(u)y(v) \rangle - \sqrt{2/\gamma} \langle y(u) \text{Im}[\nu_1(v)] + y(v) \text{Im}[\nu_1(u)] \rangle \} \cos^2(\theta - \Theta) \\ &\quad + (E')^2 [\cos(\theta - \Theta) - 1]^2 + (2/\gamma) \delta(u - v). \end{aligned} \quad (\text{B19})$$

Substituting gives

$$\begin{aligned}\langle I_x(u)I_x(v) \rangle &= e^{-r} e^{-\gamma(1+\varepsilon)|u-v|/2} \cos^2(\theta - \Theta) + \{e^r e^{-\gamma(1-\varepsilon)|u-v|/2} + (E')^2\} \sin^2(\theta - \Theta) + (2/\gamma)\delta(u-v) \\ \langle I_y(u)I_y(v) \rangle &= e^{-r} e^{-\gamma(1+\varepsilon)|u-v|/2} \sin^2(\theta - \Theta) + e^r e^{-\gamma(1-\varepsilon)|u-v|/2} \cos^2(\theta - \Theta) \\ &\quad + (E')^2 [\cos(\theta - \Theta) - 1]^2 + (2/\gamma)\delta(u-v).\end{aligned}\tag{B20}$$

Using this result gives

$$\begin{aligned}\langle (I_x - \bar{x})^2 + (I_y - \bar{y})^2 \rangle &= \text{const.} + (E')^2 \left(1 - \frac{\sigma_x^2 - 1}{\sqrt{\varepsilon^2 + 1}}\right)^2 \sin^2(\theta - \Theta) + (E')^2 \left(1 - \frac{\sigma_y^2 - 1}{\sqrt{\varepsilon^2 + 1}}\right)^2 [\cos(\theta - \Theta) - 1]^2 \\ &\quad + \left(\frac{e^r}{\sigma_x^2 + 1} - \frac{e^{-r}}{\sigma_y^2 + 1}\right) \left[\frac{(\sigma_x^2 - 1)^2 - (\sigma_y^2 - 1)^2}{\sqrt{\varepsilon^2 + 1}} - 2(\sigma_x^2 - \sigma_y^2)\right] \sin^2(\theta - \Theta).\end{aligned}\tag{B21}$$

Here the terms that do not depend on θ or Θ have been collected into the constant. Expanding to second order in $(\theta - \Theta)$ and using the equilibrium values of σ_x^2 and σ_y^2 gives

$$\frac{\gamma dt}{2} \langle (I_x - \bar{x})^2 + (I_y - \bar{y})^2 \rangle = \text{const.} + \frac{E^2}{1 + e^{-2r}} (\theta - \Theta)^2 + 2\gamma dt \left[\cosh r - (\varepsilon^2 + 1)^{-1/2} \right] (\theta - \Theta)^2.\tag{B22}$$

ACKNOWLEDGMENTS

This research has been supported by the Australian Research Council. The authors acknowledge valuable dis-

cussions with Elanor Huntington, Hideo Mabuchi, and Tim Ralph.

-
- [1] W. G. Unruh, in *Quantum Optics, Experimental Gravitation, and Measurement Theory*, edited by P. Meystre and M. O. Scully (Plenum, New York, 1982), p. 647.
 - [2] U. Leonhardt, J. A. Vaccaro, B. Böhmer, and H. Paul, *Phys. Rev. A* **51**, 84 (1995).
 - [3] G. S. Summy and D. T. Pegg, *Opt. Comm.* **77**, 75 (1990).
 - [4] K. L. Pregnell and D. T. Pegg, *Phys. Rev. Lett.* **89**, 173601 (2002).
 - [5] H. M. Wiseman, *Phys. Rev. Lett.* **75**, 4587 (1995); H. M. Wiseman and R. B. Killip, *Phys. Rev. A* **56**, 944 (1997); H. M. Wiseman and R. B. Killip, *Phys. Rev. A* **57**, 2169 (1998); D. Berry, H. M. Wiseman, and Z. X. Zhang, *Phys. Rev. A* **60**, 2458 (1999); D. W. Berry and H. M. Wiseman *Phys. Rev. A* **63**, 013813 (2001).
 - [6] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, *Phys. Rev. Lett.* **89**, 133602 (2002).
 - [7] D. W. Berry and H. M. Wiseman, *Phys. Rev. A* **65**, 043803 (2002).
 - [8] D. T. Pope, H. M. Wiseman, and N. K. Langford, *Phys. Rev. A* **70**, 043812 (2004).
 - [9] K. McKenzie, E. Mikhailov, K. Goda, P. K. Lam, N. Grosse, M. B. Gray, N. Mavalvala, and D. E. McClelland, *J. Opt. B: Quantum Semiclass. Opt.* **7**, S421 (2005).
 - [10] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer, Berlin, 2000).
 - [11] D. W. Berry, Ph.D. thesis, the University of Queensland (2001); quant-ph/0202136.
 - [12] P. D. Drummond and Z. Fizek, *Quantum Squeezing* (Springer-Verlag, Berlin 2004).
 - [13] H. M. Wiseman and G. J. Milburn, *Phys. Rev. A* **47**, 1652 (1993).
 - [14] E. Huntington, (private communication, 2005).
 - [15] A. S. Holevo, *Springer Lecture Notes in Math.* **1055**, 153 (1984).